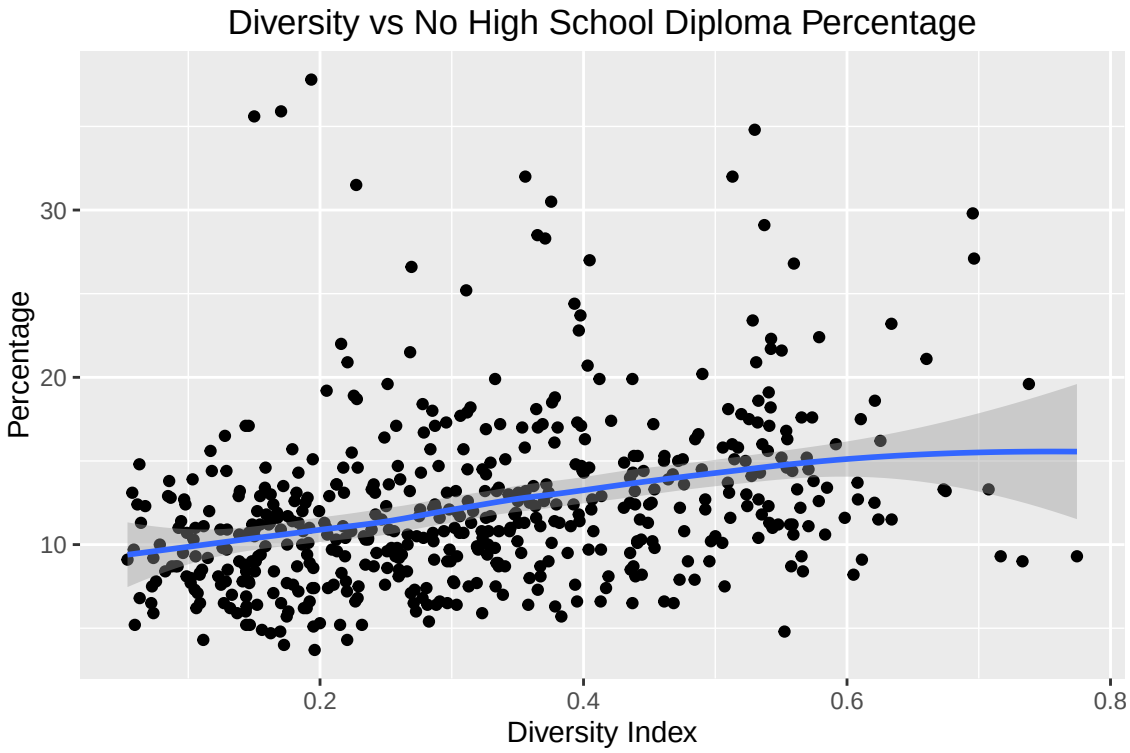
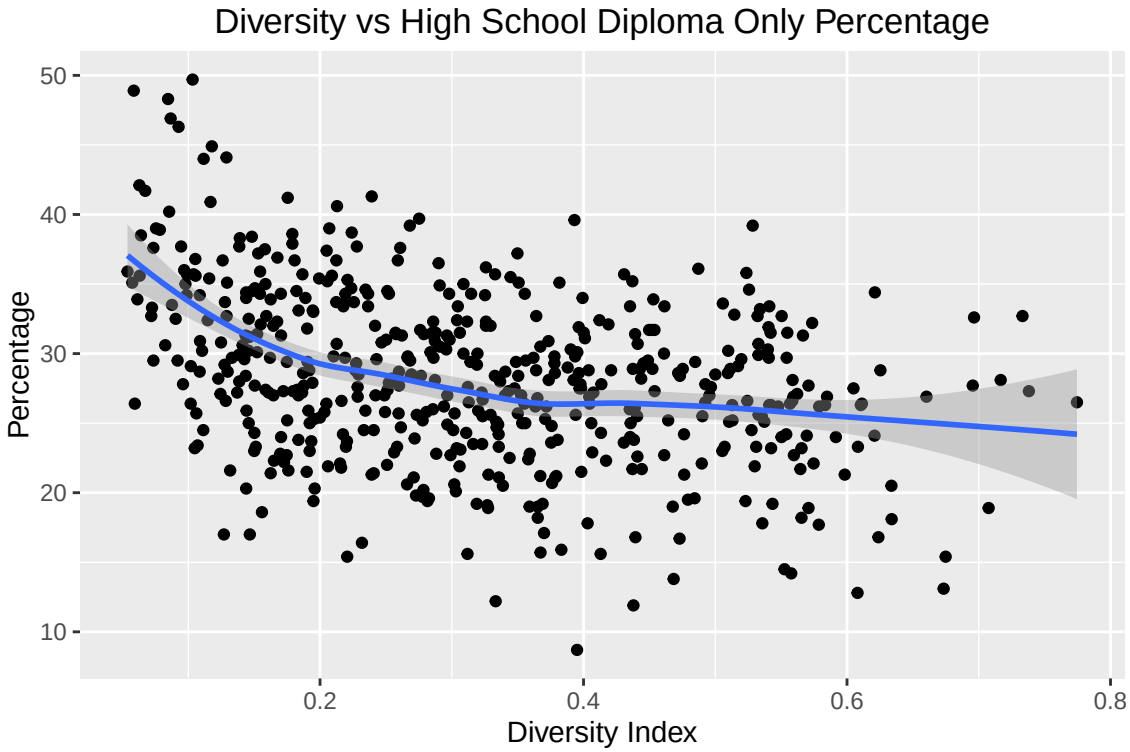
# STAT 605 Progress Report

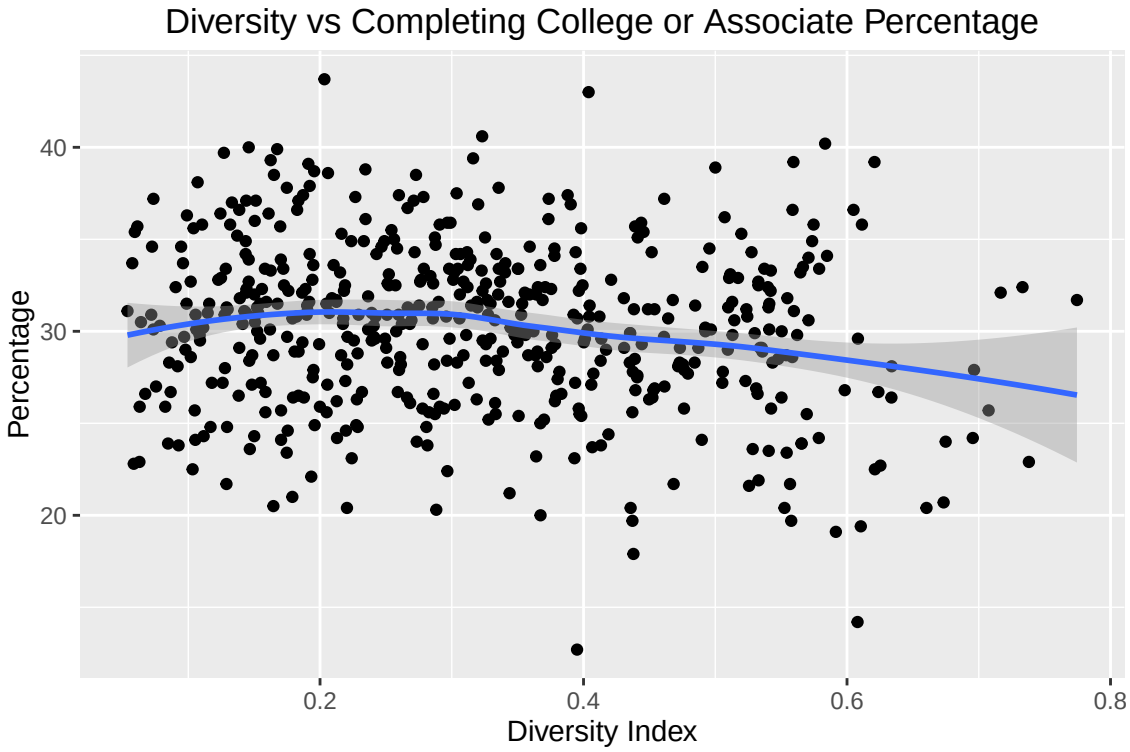
*Bing Xu*

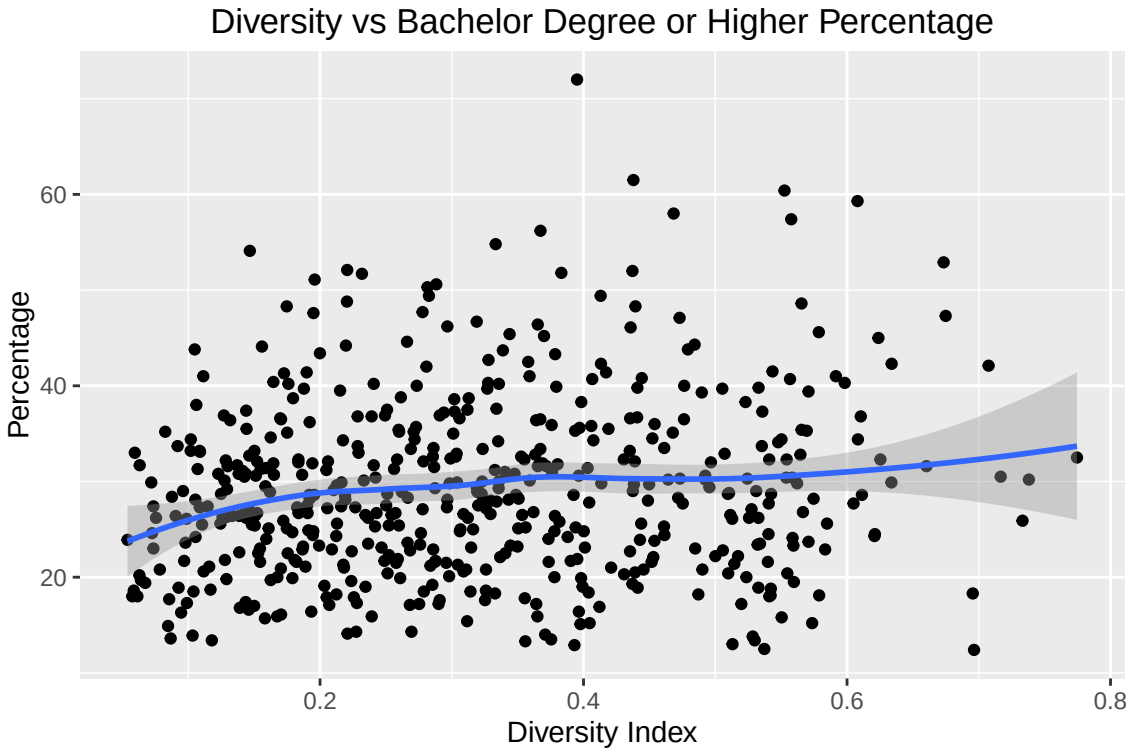
*11/18/2016*

We also find a data set that include the percentage of attaining certain education level for each county. So we are interested in finding whether there are relationships bewtween the diversity we devised and percentages of completing degrees. There are four percentages, which are percent of adults with less than a high school diploma, percent of adults with a high school diploma only, percent of adults completing some college or associate's degree, percent of adults with a bachelor's degree or higher.




Diversity vs High School Diploma Only Percentage
Percentage
50
40
30
20
10
0.2
0.4
0.6
0.8
Diversity Index


Diversity vs Completing College or Associate Percentage
Percentage
40
30
20
0.2
0.4
0.6
0.8
Diversity Index




From graphs above we can see that all rates seems not to have strong correlation with the diversity index. We also conducted the linear regression analysis and all regressions are having a relatively poor $R^2$, which the maximum is 13.27% (The regression between diversity index and the percent of completing a high school diploma only).

We can find that there is a decreasing trend between the diversity index and the percentage of adults completing a high school degree only. We can find that counties whose high school diploma only percentage is high are having a relatively lower diversity index, which means that the race are more concentrated in these counties. As for the relationship between the percentage of less than a high school diploma and the diversity index, the trend is a bit oppoiste. Highly diversified counties are a bit likely to have a higher percentage. There are some outliers and these outliers implies that counties with either low or high diversity indexes could have a high percentage.

While for the last two graphs, the relationships between variables are quite weak and there are many outliers in degree percentages.